



OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/059,447

DATE: 05/31/2002

TIME: 08:29:37

Input Set : A:\SMITH.ST25.txt

Output Set: N:\CRF3\05312002\J059447.raw

**Does Not Comply**  
**Corrected Diskette Needed**

3 <110> APPLICANT: Smith, Dan  
 5 <120> TITLE OF INVENTION: ALPHA-N-ACETYL GALACTOSAMINIDASE FROM CLOSTRIDIUM PERFRINGENS  
 7 <130> FILE REFERENCE: 0994.00134  
 9 <140> CURRENT APPLICATION NUMBER: 10/059,447  
 10 <141> CURRENT FILING DATE: 2002-01-29  
 12 <150> PRIOR APPLICATION NUMBER: 60/064,683  
 13 <151> PRIOR FILING DATE: 1997-11-03  
 15 <150> PRIOR APPLICATION NUMBER: 09/185,476  
 16 <151> PRIOR FILING DATE: 1998-11-03  
 18 <160> NUMBER OF SEQ ID NOS: 1  
 20 <170> SOFTWARE: PatentIn version 3.0

## ERRORED SEQUENCES

22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 29  
 24 <212> TYPE: PRT  
 25 <213> ORGANISM: Clostridium perfringens  
 27 <220> FEATURE:  
 W--> 28 <221> NAME/KEY: Xaa  
 29 <222> LOCATION: (1)..(29)  
 30 <223> OTHER INFORMATION: Xaa is any protein  
 33 <400> SEQUENCE: 1  
 35 Met Lys Val Leu Gly Asn Tyr Ile Gln Arg Asn Phe His Tyr Asp Gly  
 E--> 36 1 5 10 10 15 15  
 W--> 38 Lys Xaa Phe Tyr Thr Lys Gln Phe Asn Lys Pro Ile Xaa  
 39 20 25

*Xaa can only represent a single  
 amino  
 acid*

*misaligned amino acid number  
 (see item 3 on Error Summary Sheet)*

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/059,447

DATE: 05/31/2002

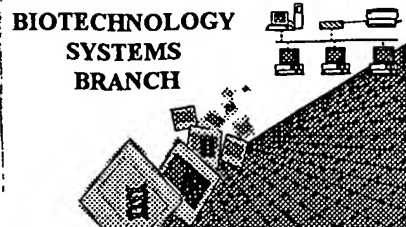
TIME: 08:29:38

Input Set : A:\SMITH.ST25.txt

Output Set: N:\CRF3\05312002\J059447.raw

L:28 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1  
L:36 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:1  
L:38 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:16

0590  
2523



## **RAW SEQUENCE LISTING** **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/059,447  
Source: O/P  
Date Processed by STIC: 5/31/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

**<http://www.uspto.gov/web/offices/pac/checker>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom!

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER: 10/059,447

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos  
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2      Invalid Line Length  
The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering  
The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII  
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length  
Sequence(s)          contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"  
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)         . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)  
Sequence(s)          missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(ii) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped  
  
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8      Skipped Sequences  
    (NEW RULES)  
Sequence(s)          missing. If intentional, please insert the following lines for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 9      Use of n's or Xaa's  
    (NEW RULES)  
Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10      Invalid <213>  
    Response  
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11      Use of <220>  
Sequence(s)          missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"  
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n  
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.